A USEFUL NEW TYPE OF RANDOM REGRESSIONS BASED ON BIOLOGICAL DIFFERENCES AMONG REPEATED RECORDS, APPLICATION TO LONGEVITY

N. Gengler[1,2], S. Vanderick[1], C. Croquet[1,2], H. Soyeurt[1,3] and P. Mayeres[1,4]

[1] Animal Science Unit, Gembloux Agricultural University, B-5030 Gembloux, Belgium
[2] National Fund for Scientific Research, B-1000 Brussels, Belgium
[3] F.R.I.A., B-1000 Brussels, Belgium
[4] Walloon Breeding Association, B-5590 Ciney, Belgium

INTRODUCTION

A major problem in random regression models is that it is not always obvious what type of regressions to use. Different types of functions were identified and use. The first type used where functions that described lactation shapes. These functions did an excellent job to describe the mean, however were very poor in modeling of (co)variance structures. The second group of functions was based on strictly mathematical ones as polynomials. Polynomials were excellent for modeling the (co)variances as long as high order polynomials could be used. Different alternative functions were proposed over time (e.g., splines). Recently another alternative method was proposed by Wiggans and Van Raden (2004) based on the concept of parity differences (PD). Instead of using predefined functions they defined as regressions differences among repeated records. This can be considered an approximation of expected a priory change in genetic merit across those repetitions where the relative size of genetic differences by parity were derived from genetic correlations. We will use the word biological differences as the idea is to base it on individual difference corrected for the environment. The following example might clarify the general idea. Wiggans and Van Raden (2004) defined relative PD among the first five lactation for milk yield as -0.9, 0.1, 0.4, 0.6 and 0.7 which means that differences from second to third, from third to fourth and from fourth to five represent 30%, 20% respectively 10% of the difference from first to second. This has the side effect that (co)variance structures are modelled as quadratic functions of regressors. Through the use of these PD they linearized changes from one lactation to the next. The objective of this paper was to present this idea and to use it for multi-lactation longevity evaluations.

MATERIAL AND METHODS

Application to longevity. There is still a certain degree of diversity in the genetic evaluation models for longevity. Roughly one can say that there are 3 groups of models used (INTERBULL, 2005): productive life or lifespan models and survival models where those models can be subdivided into linear and non-linear (survival analysis) models. Optimal use of the already known survival history of a given cows was important, allowing to model her survival at a given moment in a given herd amongst other animals in their respective contemporary groups. This favored an approach that is similar to the currently in Canada used MT model (INTERBULL, 2005) which considers survival in the first three lactations as three different traits. Australia uses a repeatability model on survival data for cows over successive years (INTERBULL, 2005). A natural extension to this type of models is a random regression model that combines both ideas modeling genetic differences across lactation survival with a computational straight forward extension to a repeatability model (Veerkamp *et al.*, 2001). This model was also chosen because initial research using other models failed totally to get reasonable breeding values for longevity.

Definition of random regressions. PD based on the proportions of survival are shown in Table 1. As these proportions are direct functions of survival probabilities they should reflect a least partially the expected change in genetic merit across parities. Also as population means

across environments they should reflect general biological differences. Using only a linear function can be in a lot of situations suboptimal and a quadratic function might be necessary in order to allow at least quadratic variation of genetic merit. A constant genetic effect was also introduced. The use of constant, linear and quadratic functions of PD however did not guarantied correct representation of genetic variances. Therefore additional steps were taken to adjust regressions during the variance components estimation procedures.

Table 1. Statistics on the Holstein data used for Walloon official evaluation for Holsteins in May 2005, together with parity difference (PD) regressions.

			Regressions		
Parity	Lactations	Proportion	PD^0	PD^1	PD^2
1	392890	0.348	1	0.28	0.0795
2	274596	0.243	1	0.18	0.0313
3	186588	0.165	1	0.10	0.0098
4	121808	0.108	1	0.04	0.0018
5	74496	0.066	1	0.00	0.0000
6	41691	0.037	1	-0.03	0.0008
7	21585	0.019	1	-0.05	0.0022
8	10022	0.009	1	-0.06	0.0032
9	4315	0.004	1	-0.06	0.0038
10	1611	0.001	1	-0.07	0.0042
11	603	0.001	1	-0.07	0.0042
12	213	0.000	1	-0.07	0.0044
13	74	0.000	1	-0.07	0.0044
14	27	0.000	1	-0.07	0.0044
15	10	0.000	1	-0.07	0.0044
16	3	0.000	1	-0.07	0.0044
17	1	0.000	1	-0.07	0.0044

Genetic evaluation model. The general evaluation model can be written as $\mathbf{y = Xh + Wb + ZQa + e}$ where **y** is a vector, h is a vector of fixed contemporary effects based of herd x quota-year (1^{st} april to 31^{st} of march) of calving x parity-group (1, 2, 3, 4 and +), **b** is a vector of fixed effects of birth-year x parity-group (1, 2, 3, 4 and +), a is a vector of the three random regression effects per animal, **e** is vector of the random residual effects, **X**, **W** **Z** are incidence matrices linking observations to the effects and **Q** is a matrix of regressors. This modelled allowed correlation among parities to vary. The trend correction effect (**b**) plays an important role because it allowed us to take a least partially into account the trend due to selection on correlated traits. For the moment we prefer this approach over that of direct regression on phenotypes because it is still unclear what the reasons for voluntary culling in our population are.

Estimation of genetic parameters. A major challenge was the estimation of variance components. They were developed from various sources. Initial estimations of correlations among first three parities were taken from literature (e.g., Jarirath *et al.*, 1998). Based on a subset of data correlations and variances across the first five parities were obtained. Heritability of survival in a given parity was put to 0.03 for all parities. Genetic (co)variances among initial regressions were obtained by backsolving. Regressions were transformed and standardized by adjusting regressors and therefore variances towards the variance of the fifth parity. Residuals were assumed uncorrelated and residuals variances standardised towards the fifth lactation using weights. The assumption of uncorrelated residuals or no permanent

environmental effect was based on own preliminary research and on the theoretically differences to explain non-genetic animal effects across lactations.

RESULTS AND DISCUSSION

Final, uncorrelated regressions and weights used in the computations are in Table 2. Using these regressions allowed defining uncorrelated random regression effects and correcting variances across all lactations.

Table 2. Final uncorrelated regressions and weights used in the computations.

	Parity										
Regression	1	2	3	4	5	6	7	8	9	10	11-17
1	0.42	0.19	0.50	0.81	1.00	0.99	0.98	0.98	0.98	0.98	0.97
2	0.41	0.71	0.59	0.24	0.00	-0.13	-0.18	-0.21	-0.22	-0.22	-0.23
3	0.15	0.07	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Weight	2.80	1.83	1.68	1.41	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 3 gives details on the variances and correlations obtained. Results for the second lactation are interesting. This lactation behaves differently from the others. The results in Table 3 show also that the random regressions defined on PD were able to describe the high dimensional (17 x 17) (co)variances needed for an equivalent multi-lactation model. National evaluations were based on 1,130,533 lactation records from 392,890 Holstein (at least 75% Holstein or Red-Holstein genes) cows. Genetic evaluations were submitted to INTERBULL for 1,394 sires and a total of 436 sires were used by INTERBULL in the routine run. Statistics comparing the correlations of our evaluation with the other INTERBULL populations are given in Table 4. Results indicated that our breeding values behaved well and had reasonably high correlations with most other populations (average 0.69).

Table 3. Genetic correlations, genetic variances (on diagonal) and residual variances (residual correlations being 0) across selected lactations (1 to 7, 10 and 17).

Parity	1	2	3	4	5	6	7	10	17
	Genetic								
1	**0.00232**	0.850	0.972	0.858	0.694	0.604	0.561	0.525	0.523
2	0.850	**0.00354**	0.900	0.514	0.251	0.129	0.074	0.029	0.026
3	0.872	0.900	**0.00386**	0.835	0.645	0.545	0.497	0.458	0.455
4	0.858	0.514	0.835	**0.00496**	0.959	0.917	0.893	0.871	0.870
5	0.694	0.251	0.645	0.959	**0.00650**	0.992	0.984	0.975	0.974
6	0.604	0.129	0.545	0.917	0.992	**0.00650**	0.998	0.995	0.995
7	0.561	0.074	0.497	0.893	0.984	0.998	**0.00650**	0.999	0.999
10	0.525	0.029	0.458	0.871	0.975	0.995	0.999	**0.00650**	0.999
17	0.523	0.026	0.455	0.870	0.974	0.995	0.999	0.999	**0.00650**
	Residual								
	0.0751	**0.115**	**0.125**	**0.149**	**0.210**	**0.210**	**0.210**	**0.210**	**0.210**

Table 4. Mean (of the Walloon Region), overall mean of all populations, minimum and maximum correlations used in the May 2005 routine-run.

Belgium (Walloon Region)	Overall	Minimum	Maximum
0.69	0.63	0.33	0.86

CONCLUSIONS

A new type of random regressions and its application to longevity was presented. It uses biological differences among repeated records as an approximation of expected a priory change in genetic merit across repetitions and as a way to model (co)variances. It allowed the development of a genetic evaluation system of cow survival using a lactation random regression model equivalent to a 17 trait lactation survival model. This model is currently used in the Walloon Region of Belgium for routine genetic evaluation.

REFERENCES

INTERBULL. 2005. *http://www-interbull.slu.se/national_ges_info2/framesida-ges.htm* (Accessed May 31, 2005).

Jairath L., Dekkers, J.C.M., Schaeffer, L.R., Liu, Z. Burnside, E.B. and Kolstad, B. 1998. *J. Dairy Sci.* **81**: 550-562.

Wiggans, G.R. and VanRaden P.M. 2004. *J. Dairy Sci*. **87**(Suppl. 1): 412 (abstr. 737).

Veerkamp,R.F., Brotherstone, S., Engel, B. and Meuwissen, T.H.E. 2001. Anim. Sci. **72**: 1-10.